

(12)

(21) 2 331 641

(22) 05.05.1999

(51) Int. Cl. 6: **C07K 16/00, C07K 16/28,  
A61K 39/395, G01N 33/53,  
C12N 15/63**

(85) 03.11.2000

(86) PCT/DE99/01350

(87) WO99/57150

(30) 198 19 846.9 DE 05.05.1998

(71) **DEUTSCHES KREBSFORSCHUNGSZENTRUM  
STIFTUNG DES <sup>TM</sup>FFENTLICHEN RECHTS,  
Im Neuenheimer Feld 280, HEIDELBERG, XX (DE).**

(72) **LITTLE, MELVYN (DE).  
KIPRIYANOV, SERGEJ (DE).**

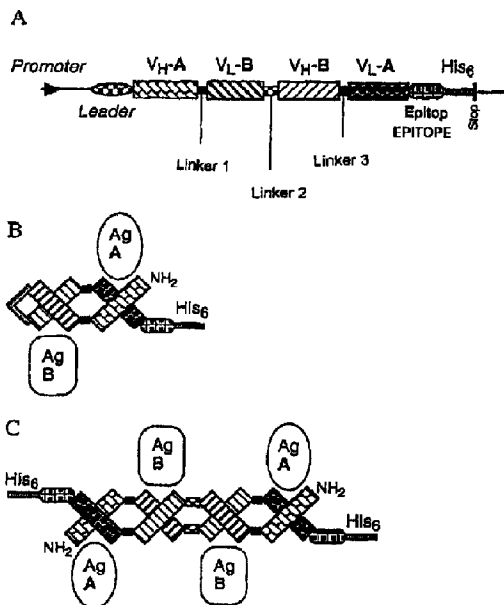
(74) **BERESKIN & PARR**

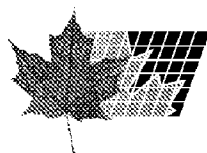
(54) CONSTRUCTIONS D'ANTICORPS MULTIVALENTES

(54) MULTIVALENT ANTIBODY CONSTRUCTS

(57)

The invention relates to a multivalent Fv antibody construct comprising at least four variable domains which are connected to one another via peptide linkers 1, 2 and 3. The invention also relates to expression plasmids which code for such an Fv antibody construct. In addition, the invention relates to a method for producing the Fv antibody constructs and to the use thereof.





(72) LITTLE, MELVYN, DE

(72) KIPRIYANOV, SERGEJ, DE

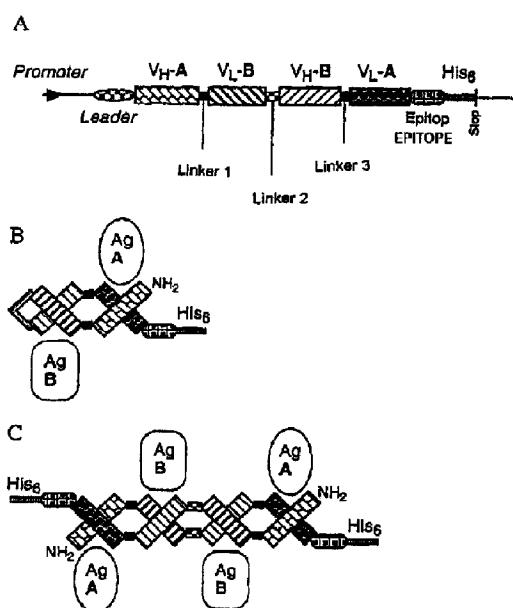
(71) DEUTSCHES KREBSFORSCHUNGSZENTRUM STIFTUNG DES  
ÖFFENTLICHEN RECHTS, DE

(51) Int.Cl.<sup>6</sup> C07K 16/00, C12N 15/63, G01N 33/53, A61K 39/395,  
C07K 16/28

(30) 1998/05/05 (198 19 846.9) DE

(54) **CONSTRUCTIONS D'ANTICORPS MULTIVALENTES**

(54) **MULTIVALENT ANTIBODY CONSTRUCTS**



(57) La présente invention concerne une construction d'anticorps  $F_v$  multivalente, comportant au moins quatre domaines variables qui sont reliés l'un à l'autre par l'intermédiaire des segments peptidiques 1, 2 et 3. L'invention concerne en outre des plasmides d'expression qui codent pour une telle construction d'anticorps  $F_v$ , ainsi qu'un procédé de réalisation des constructions d'anticorps  $F_v$  et leur utilisation.

(57) The invention relates to a multivalent  $F_v$  antibody construct comprising at least four variable domains which are connected to one another via peptide linkers 1, 2 and 3. The invention also relates to expression plasmids which code for such an  $F_v$  antibody construct. In addition, the invention relates to a method for producing the  $F_v$  antibody constructs and to the use thereof.





**PCT**  
WELTORGANISATION FÜR GEISTIGES EIGENTUM  
Internationales Büro  
INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE  
INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

<p>(51) Internationale Patentklassifikation <sup>6</sup> : <b>C07K 16/00</b></p>	<p><b>A2</b></p>	<p>(11) Internationale Veröffentlichungsnummer: <b>WO 99/57150</b></p> <p>(43) Internationales Veröffentlichungsdatum: 11. November 1999 (11.11.99)</p>
<p>(21) Internationales Aktenzeichen: PCT/DE99/01350</p> <p>(22) Internationales Anmeldedatum: 5. Mai 1999 (05.05.99)</p> <p>(30) Prioritätsdaten: 198 19 846.9      5. Mai 1998 (05.05.98)      DE</p> <p>(71) Anmelder (für alle Bestimmungsstaaten ausser US): DEUTSCHES KREBSFORSCHUNGSZENTRUM STIFTUNG DES ÖFFENTLICHEN RECHTS [DE/DE]; Im Neuenheimer Feld 280, D-69120 Heidelberg (DE).</p> <p>(72) Erfinder; und</p> <p>(75) Erfinder/Anmelder (nur für US): LITTLE, Melvyn [GB/DE]; Fritz-von-Briesen-Strasse 10, D-69151 Neekargemünd (DE). KIPRIYANOV, Sergej [RU/DE]; Furtwänglerstrasse 3, D-69121 Heidelberg (DE).</p> <p>(74) Anwalt: HUBER, Bernard; Huber &amp; Schlüssler, Truderinger Strasse 246, D-81825 München (DE).</p>		
<p>(81) Bestimmungsstaaten: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO Patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</p> <p>Veröffentlicht Ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts.</p>		
<p>(54) Title: MULTIVALENT ANTIBODY CONSTRUCTS</p> <p>(54) Bezeichnung: MULTIVALENTE ANTIKÖRPER-KONSTRUKTE</p> <p>(57) Abstract</p> <p>The invention relates to a multivalent F<sub>v</sub> antibody construct comprising at least four variable domains which are connected to one another via peptide linkers 1, 2 and 3. The invention also relates to expression plasmids which code for such an F<sub>v</sub> antibody construct. In addition, the invention relates to a method for producing the F<sub>v</sub> antibody constructs and to the use thereof.</p> <p>(57) Zusammenfassung</p> <p>Die vorliegende Erfindung betrifft ein multivalentes F<sub>v</sub>-Antikörper-Konstrukt mit mindestens vier variablen Domänen, die über die Peptidlinker 1, 2 und 3 miteinander verbunden sind. Ferner betrifft die Erfindung Expressionsplasmide, die für ein solches F<sub>v</sub>-Antikörper-Konstrukt codieren, und ein Verfahren zur Herstellung der F<sub>v</sub>-Antikörper-Konstrukte sowie deren Verwendung.</p>		
<div style="display: flex; justify-content: space-around;"> <div style="width: 45%;"> <p><b>A</b></p> </div> <div style="width: 45%;"> <p><b>B</b></p> </div> <div style="width: 45%;"> <p><b>C</b></p> </div> </div>		

Applicant: Deutsches Krebsforschungszentrum  
Attorney's File: K 2675

### **Multivalent Antibody Constructs**

The present invention relates to multivalent  $F_v$  antibody constructs, expression plasmids which code for them, and a method for producing the  $F_v$  antibody constructs as well as the use thereof.

Natural antibodies are dimers and are therefore referred to as bivalent. They have four variable domains, namely two  $V_H$  domains and two  $V_L$  domains. The variable domains serve as binding sites for an antigen, a binding site being formed from a  $V_H$  domain and a  $V_L$  domain. Natural antibodies recognize one antigen each, so that they are also referred to as monospecific. Furthermore, they also have constant domains which add to the stability of the natural antibodies. On the other hand, they are also co-responsible for undesired immune responses which result when natural antibodies of various animal species are administered mutually.

In order to avoid such immune responses, antibodies are constructed which lack the constant domains. In particular, these are antibodies which only comprise the variable domains. Such antibodies are designated  $F_v$  antibody constructs. They are often available in the form of single-chain monomers paired with one another.

However, it showed that  $F_v$  antibody constructs only have little stability. Therefore, their usability for therapeutic purposes is strongly limited.

Thus, it is the object of the present invention to provide an antibody by means of which undesired immune responses can be avoided. Furthermore, it shall have a stability which makes it usable for therapeutic uses.

According to the invention this is achieved by the subject matters defined in the claims.

Therefore, the subject matter of the present invention relates to a multivalent  $F_v$  antibody construct which has great stability. Such a construct is suitable for diagnostic and therapeutic purposes.

The present invention is based on the applicant's insights that the stability of an  $F_v$  antibody construct can be increased if it is present in the form of a single-chain dimer where the four variable domains are linked with one another via three peptide linkers. The applicant also recognized that the  $F_v$  antibody construct folds with itself when the middle peptide linker has a length of about 10 to 30 amino acids. The applicant also recognized that the  $F_v$  antibody construct folds with other  $F_v$  antibody constructs when the middle peptide linker has a length of about up to 10 amino acids so as to obtain a multimeric, i.e. multivalent,  $F_v$  antibody construct. The applicant also realized that the  $F_v$  antibody construct can be multi-specific.

According to the invention the applicant's insights are utilized to provide a multi-valent  $F_v$  antibody construct

which comprises at least four variable domains which are linked with one another via peptide linkers 1, 2 and 3.

The expression "F<sub>v</sub> antibody construct" refers to an antibody which has variable domains but no constant domains.

The expression "multivalent F<sub>v</sub> antibody construct" refers to an F<sub>v</sub> antibody which has several, but at least four, variable domains. This is achieved when the single-chain F<sub>v</sub> antibody construct folds with itself so as to give four variable domains, or folds with other single-chain F<sub>v</sub> antibody constructs. In the latter case, an F<sub>v</sub> antibody construct is given which has 8, 12, 16, etc., variable domains. It is favorable for the F<sub>v</sub> antibody construct to have four or eight variable domains, i.e. it is bivalent or tetravalent (cf. Fig. 1). Furthermore, the variable domains may be equal or differ from one another, so that the antibody construct recognizes one or several antigens. The antibody construct preferably recognizes one or two antigens, i.e. it is monospecific and bispecific, respectively. Examples of such antigens are proteins CD19 and CD3.

The expression "peptide linkers 1, 3" refers to a peptide linker adapted to link variable domains of an F<sub>v</sub> antibody construct with one another. The peptide linker may contain any amino acids, the amino acids glycine (G), serine (S) and proline (P) being preferred. The peptide linkers 1 and 3 may be equal or differ from each other. Furthermore, the peptide linker may have a length of about 0 to 10 amino acids. In the former case, the peptide linker is only a peptide bond from the COOH residue of one of the variable domains and the NH<sub>2</sub> residue of another of the variable domains. The peptide linker preferably comprises the amino acid sequence GG.

The expression "peptide linker 2" refers to a peptide linker adapted to link variable domains of an  $F_v$  antibody construct with one another. The peptide linker may contain any amino acids, the amino acids glycine (G), serine (S) and proline (P) being preferred. The peptide linker may also have a length of about 3 to 10 amino acids, in particular 5 amino acids, and most particularly the amino acid sequence GGPGS, which serves for achieving that the single-chain  $F_v$  antibody construct folds with other single-chain  $F_v$  antibody constructs. The peptide linker can also have a length of about 11 to 20 amino acids, in particular 15 to 20 amino acids, and most particularly the amino acid sequence  $(G_4S)_4$ , which serves for achieving that the single-chain  $F_v$  antibody construct folds with itself.

An  $F_v$  antibody construct according to the invention can be produced by common methods. A method is favorable in which DNAs coding for the peptide linkers 1, 2 and 3 are ligated with DNAs coding for the four variable domains of an  $F_v$  antibody construct such that the peptide linkers link the variable domains with one another and the resulting DNA molecule is expressed in an expression plasmid. Reference is made to Examples 1 to 6. As to the expressions " $F_v$  antibody construct" and "peptide linker" reference is made to the above explanations and, by way of supplement, to Maniatis, T. et al., Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory 1982.

DNAs which code for an  $F_v$  antibody construct according to the invention also represent a subject matter of the present invention. Furthermore, expression plasmids which contain such DNAs also represent a subject matter of the present invention. Preferred expression plasmids are pDISC3x19-LL,

pDISC3x19-SL, pPIC-DISC-LL, pPIC-DISC-SL, pDISC5-LL and pDISC6-SL. The first four were deposited with the DSMZ (*Deutsche Sammlung für Mikroorganismen und Zellen*) [German-type collection for micro-organisms and cells] on April 30, 1998 under DSM 12150, DSM 12149, DSM 12152 and DSM 12151, respectively.

Another subject matter of the present invention relates to a kit, comprising:

- (a) an F<sub>v</sub> antibody construct according to the invention, and/or
- (b) an expression plasmid according to the invention, and
- (c) conventional auxiliary agents, such as buffers, solvents and controls.

One or several representatives of the individual components may be present.

The present invention provides a multivalent F<sub>v</sub> antibody construct where the variable domains are linked with one another via peptide linkers. Such an antibody construct distinguishes itself in that it contains no parts which can lead to undesired immune reactions. Furthermore, it has great stability. It also enables to bind several antigens simultaneously. Therefore, the F<sub>v</sub> antibody construct according to the invention is perfectly adapted to be used not only for diagnostic but also for therapeutic purposes. Such purposes can be seen as regards any disease, in particular a viral, bacterial or tumoral disease.



**Brief description of the drawings:**

**Fig. 1** shows the genetic organization of an F<sub>v</sub> antibody construct (A) according to the invention and schemes for forming a bivalent (B) or tetravalent F<sub>v</sub> antibody construct (C). Ag: antigen; His<sub>6</sub>: six C-terminal histidine residues; stop: stop codon (TAA); V<sub>H</sub> and V<sub>L</sub>: variable region of the heavy and light chains.

**Fig. 2** shows the scheme for the construction of the plasmids pDISC3x19-LL and pDISC3x19-SL. c-myc: sequence coding for an epitope which is recognized by the antibody 9E1, His<sub>6</sub>: sequence which codes for six C-terminal histidine residues; PelB: signal peptide sequence of the bacterial pectate lyase (PelB leader); rbs: ribosome binding site; Stop: stop codon (TAA); V<sub>H</sub> and V<sub>L</sub>: variable region of the heavy and light chains.

**Fig. 3** shows a diagram of the expression plasmid pDISC3x19-LL. 6xHis: sequence which codes for six C-terminal histidine residues; bla: gene which codes for  $\beta$ -lactamase responsible for ampicillin resistance; bp: base pairs; c-myc: sequence coding for an epitope which is recognized by the 9E10 antibody; ColE1: origin of the DNA replication; f1-IG: intergenic region of the bacteriophage f1; Lac P/O: wt lac-operon promoter/operator; linker 1: sequence which codes for a GlyGly dipeptide linking the V<sub>H</sub> and V<sub>L</sub> domains; linker 2: sequence coding for a (Gly<sub>4</sub>Ser)<sub>4</sub> polypeptide which links the hybrid scFv fragments; Pel-B leader: signal peptide sequence of the bacterial pectate lyase; rbs: ribosome binding site; V<sub>H</sub> and V<sub>L</sub>: variable region of the heavy and light chains.

**Fig. 4** shows a diagram of the expression plasmid pDISC3x19-SL. 6xHis: sequence which codes for six C-terminal histidine

residues; bla: gene which codes for  $\beta$ -lactamase which is responsible for the ampicillin resistance; bp: base pairs; c-myc: sequence coding for an epitope recognized by the 9E10 antibody; ColE1: origin of DNA replication; fl-IG: intergenic region of the bacteriophage fl; Lac P/O: wt lac-operon promoter/operator; linker 1: sequence which codes for a GlyGly dipeptide which links the  $V_H$  and  $V_L$  domains; linker 3: sequence which codes for a GlyGlyProGlySer oligopeptide which links the hybrid scFv fragments; Pel-B leader: signal peptide sequence of the bacterial pectate lyase; rbs: ribosome binding site;  $V_H$  and  $V_L$ : variable region of the heavy and light chains.

**Fig. 5** shows the nucleotide sequence and the amino acid sequence derived therefrom of the bivalent  $F_v$  antibody construct encoded by the expression plasmid pDIS3x19-LL. c-myc epitope: sequence coding for an epitope which is recognized by the antibody 9E10; CDR: region determining the complementarity; framework: framework region; His6 tail: sequence which codes for six C-terminal histidine residues; PelB leader: signal peptide sequence of the bacterial pectate lyase; RBS: ribosome binding site;  $V_H$  and  $V_L$ : variable region of the heavy and light chains.

**Fig. 6** shows the nucleotide sequence and the derived amino acid sequence of the tetravalent  $F_v$  antibody construct encoded by the expression plasmid pDISC3x19-SL. c-myc epitope: sequence coding for an epitope which is recognized by the 9E10 antibody; CDR: region determining complementarity; framework: framework region; His6 tail: sequence coding for the six C-terminal histidine residues; PelB leader: signal peptide sequence of the bacterial pectate lyase; RBS: ribosome binding site;  $V_H$  and  $V_L$ : variable region of the heavy and light chains.

**Fig. 7** shows the nucleotide sequence and the derived amino acid sequence of a connection between a gene which codes for an  $\alpha$ -factor leader sequence and a gene coding for the tetravalent  $F_v$  antibody construct in the *Pichia* expression plasmid pPIC-DISC-SL. Alpha-factor signal: leader peptide sequence of the *Saccharomyces cerevisiae*- $\alpha$  factor secretion signal;  $V_H$ : variable region of the heavy chain. Rhombs indicate the signal cleaving sites.

**Fig. 8** shows the nucleotide sequence and the derived amino acid sequence of a connection between a gene coding for an  $\alpha$ -factor leader sequence and a gene which codes for the bivalent  $F_v$  antibody construct in the *Pichia* expression plasmid pPIC-DISC-LL. Alpha-factor signal: leader peptide sequence of the *Saccharomyces cerevisiae*- $\alpha$  factor secretion signal;  $V_H$ : variable region of the heavy chain. Rhombs show the signal cleaving sites.

**Fig. 9** shows a diagram of the expression plasmid pDISC5-LL. 6xHis: sequence coding for six C-terminal histidine residues; bla: gene which codes for  $\beta$ -lactamase responsible for ampicillin resistance; bp: base pairs; c-myc: sequence coding for an epitope which is recognized by the 9E10 antibody; hok-sok: plasmid-stabilizing DNA locus; LacI: gene which codes for the Lac repressor; Lac P/O: wt lac-operon-promoter/operator; LacZ': gene which codes for the  $\alpha$ -peptide of  $\beta$ -galactosidase; linker 1: sequence which codes for a GlyGly dipeptide connecting the  $V_H$  and  $V_L$  domains; linker 2: sequence which codes for a (Gly<sub>4</sub>Ser)<sub>4</sub> polypeptide linking the hybrid scFv fragments; M13 IG: intergenic region of the M13 bacteriophage; pBR322ori: origin of DNA replication; Pel-B leader: signal peptide sequence of the bacterial pectate lyase; rbs: ribosome binding site which originates

from the *E. coli* lacZ gene (lacZ), from the bacteriophage T7 gene 10 (T7g10) or from the *E. coli* skp gene (skp); skp: gene which codes for the bacterial periplasmic factor Skp/OmpH; tHP: strong transcription terminator; tIPP: transcription terminator; V<sub>H</sub> and V<sub>L</sub>: variable region of the heavy and light chains.

**Fig. 10** shows a diagram of the expression plasmid pDISC6-SL. 6xHis: sequence which codes for six C-terminal histidine residues; bla: gene which codes for  $\beta$ -lactamase responsible for ampicillin resistance; bp: base pairs; c-myc: sequence coding for an epitope which is recognized by the 9E10 antibody; hok-sok: plasmid-stabilized DNA locus; LacI: gene which codes for the Lac repressor; Lac P/O: wt lac-operon promoter/operator; LacZ': gene which codes for the  $\alpha$ -peptide of  $\beta$ -galactosidase; linker 1: sequence which codes for a GlyGly dipeptide which links the V<sub>H</sub> and V<sub>L</sub> domains; linker 3: sequence which codes for a GlyGlyProGlySer oligopeptide linking the hybrid scFv fragments; M13 IG: intergenic region of the M13 bacteriophage; pBR322ori: origin of DNA replication; Pel-B leader: signal peptide sequence of the bacterial pectate lyase; rbs: ribosome binding site originating from the *E. coli* lacZ gene (lacZ), from the bacteriophage T7 gene 10 (T7g10) or from the *E. coli* skp gene (skp); skp: gene which codes for the bacterial periplasmic factor Skp/OmpH; tHP: strong transcription terminator; tIPP: transcription terminator; V<sub>H</sub> and V<sub>L</sub>: variable region of the heavy and light chains.

The invention is explained by the below examples.

**Example 1: Construction of the plasmids pDISC3x19-LL and pDISC3x19-SL for the expression of bivalent, bispecific and/or tetravalent, bispecific F<sub>v</sub> antibody constructs in bacteria**

The plasmids pHOG- $\alpha$ CD19 and pHOG-dmOKT3 which code for the scFv fragments derived from the hybridoma HD37 which is specific to human CD19 (Kipriyanov et al., 1996, J.-Immunol. Meth. 196, 51-62) and from the hybridoma OKT3 which is specific to human CD3 (Kipriyanov et al., 1997, Protein Eng. 10, 445-453), respectively, were used for the construction of expression plasmids for a single-chain F<sub>v</sub> antibody construct. A PCR fragment 1 of the V<sub>H</sub> domain of anti-CD19, followed by a segment which codes for a GlyGly linker, was produced using the primers DP1, 5'-TCACACAGAATTC-TTAGATCTATTAAAGAGGAGAAATTAACC, and DP2, 5'-AGCACACGATATCACCGCCAAGCTTGGGTGTTGTTTGGC (cf. Fig. 2). The PCR fragment 1 was cleaved by EcoRI and EcoRV and ligated with the EcoRI/EcoRV-linearized plasmid pHOG-dmOKT3 so as to produce the vector pHOG19-3. The PCR fragment 2 of the V<sub>L</sub> domain of anti-CD19, followed by a segment which codes for a c-myc epitope and a hexahistidinyI tail, was produced using the primers DP3, 5'-AGCACACAAGCTTGGCGGTGATATCTTGCTCACCCAAAC-TCCA, and DP4, 5'-AGCACACTCTAGAGACACACAGATCTTTAGTGATGGTGATGGTGATGTGAGTTTAGG. The PCR fragment 2 was cleaved by HindIII and XbaI and ligated with the HindIII/XbaI-linearized plasmid pHOG-dmOKT3 so as to obtain the vector pHOG3-19 (cf. Fig. 2). The gene coding for the hybrid scFv-3-19 in the plasmid pHOG3-19 was amplified by means of PCR with the primers Bi3sk, 5'-CAGCCGGCCATGGCGCAGGTGCAACTGCAGCAG and either Li-1, 5'-TATATACTGCAGCTGCACCTGGCTACCACCACCACCGGAGCCG-CCACCACCGCTACCACCGCCGCCAGAACCACCACCACCAGCGGCCGCAGCATCAGCCCG, for the production of a long flexible (Gly<sub>4</sub>Ser)<sub>4</sub> inter-scFv linker (PCR fragment 3, cf. Fig. 2) or Li-2, 5'-TATATA-

CTGCAGCTGCACCTGCGACCCTGGGGCCACCAGCGGCCGCAGCATCAGCCCG, for the production of a short rigid GPGS linker (PCR fragment 4, cf. Fig. 2). The expression plasmids pDISC3x19-LL and pDISC3x19-SL were constructed by ligating the NcoI/PvuII restriction fragment from pHOG19-3, comprising the vector framework and the NcoI/PvuII-cleaved PCR fragments 3 and 4, respectively (cf. Figs. 3, 4). The complete nucleotide and protein sequences of the bivalent and tetravalent F<sub>v</sub> antibody constructs are indicated in Figs 5 and 6, respectively.

**Example 2: Construction of the plasmids pPIC-DISC-LL and pPIC-DISC-SL for the expression of bivalent, bispecific and/or tetravalent, bispecific F<sub>v</sub> antibody constructs in yeast**

(A) Construction of pPIC-DISC-SL

The vector pPICZαA (Invitrogen BV, Leek, Netherlands) for the expression and secretion of recombinant proteins in the yeast *Pichia pastoris* was used as a starting material. It contains a gene which codes for the *Saccharomyces cerevisiae* α-factor secretion signal, followed by a polylinker. The secretion of this vector is based on the dominant selectable marker, Zeocin<sup>TM</sup> which is bifunctional in both *Pichia* and *E. coli*. The gene which codes for the tetravalent F<sub>v</sub> antibody construct (scDia-SL) was amplified by means of PCR by the template pDISC3x19-SL using the primers 5-PIC, 5'-CCGTGAATTCCAGGTGCAACTGCAGCAGTCTGGGGCTGAACTGGC, and pSEXBn 5'-GGTCGACGTTAACCGACAAACAACAGATAAAACG. The resulting PCR product was cleaved by EcoRI and XbaI and ligated in EcoRI/XbaI-linearized pPICZαA. The expression plasmid pPIC-DISC-SL was obtained. The nucleotide and protein sequences

of the tetravalent F<sub>v</sub> antibody construct are shown in Fig. 7.

(B) Construction of pPIC-DISC-LL

The construction of pPIC-DISC-LL was carried out on the basis of pPICZαA (Invitrogen BV, Leek, Netherlands) and pDISC3x19-LL (cf. Fig. 3). The plasmid-DNA pPICZαA was cleaved by EcoRI. The overhanging 5'-ends were filled using a Klenow fragment of the *E. coli* DNA polymerase I. The resulting DNA was cleaved by XbaI, and the large fragment comprising the pPIC vector was isolated. Analogous thereto the DNA of pDISC3x19-LL was cleaved by NcoI and treated with a Klenow fragment. Following the cleavage using XbaI a small fragment, comprising a gene coding for the bivalent F<sub>v</sub> antibody, was isolated. Its ligation with a pPIC-derived vector-DNA resulted in the plasmid pPIC-DISC-LL. The nucleotide and protein sequences of the bivalent F<sub>v</sub> antibody construct are shown in Fig. 8.

**Example 3: Expression of the tetravalent and/or bivalent F<sub>v</sub> antibody construct in bacteria**

*E. coli* XL1-blue cells (Stratagene, La Jolla, CA) which had been transformed with the expression plasmids pDISC3x19-LL and pDISC3x19-SL, respectively, were cultured overnight in 2xYT medium with 50 µg/ml ampicillin and 100 mM glucose (2xYT<sub>Ga</sub>) at 37°C. 1:50 dilutions of the overnight cultures in 2xYT<sub>Ga</sub> were cultured as flask cultures at 37°C while shaking with 200 rpm. When the cultures had reached an OD<sub>600</sub> value of 0.8, the bacteria were pelleted by 10-minute centrifugation with 1500 g at 20°C and resuspended in the same volume of a fresh 2xYT medium containing 50 µg/ml ampicillin and 0.4 M saccharose. IPTG was added up to a

final concentration of 0.1 mM, and the growth was continued at room temperature (20-22°C) for 18 - 20 h. The cells were harvested by 10-minute centrifugation with 5000 g at 4°C. The culture supernatant was held back and stored on ice. In order to isolate the soluble periplasmic proteins, the pelleted bacteria were resuspended in 5 % of the initial volume of ice-cold 50 mM Tris-HCl, 20 % saccharose, 1 mM EDTA, pH 8.0. Following 1 hour of incubation on ice with occasional stirring the spheroplasts were centrifuged with 30,000 g at 4°C for 30 minutes, the soluble periplasmic extract being obtained as supernatant and the spheroplasts with the insoluble periplasmic material being obtained as pellet. The culture supernatant and the soluble periplasmic extract were combined and clarified by further centrifugation (30,000 g, 4°C, 40 min.). The recombinant product was concentrated by ammonium sulfate precipitation (final concentration 70 % saturation). The protein precipitate was obtained by centrifugation (10,000 g, 4°C, 40 min.) and dissolved in 10 % of the initial volume of 50 mM Tris-HCl, 1 M NaCl, pH 7.0. An immobilized metal affinity chromatography (IMAC) was carried out at 4°C using a 5 ml column of chelating sepharose (Pharmacia) which was charged with  $\text{Cu}^{2+}$  and had been equilibrated with 50 mM Tris-HCl, 1 M NaCl, pH 7.0 (starting buffer). The sample was loaded by passing it over the column. It was then washed with twenty column volumes of starting buffer, followed by starting buffer with 50 mM imidazole until the absorption at 280 nm of the effluent was at a minimum (about thirty column volumes). The absorbed material was eluted with 50 mM Tris-HCl, 1 M NaCl, 250 mM imidazole, pH 7.0.

The protein concentrations were determined with the Bradford dye binding test (1976, Anal. Biochem. 72, 248-254) using the Bio-Rad (Munich, Germany) protein assay kit. The



concentrations of the purified tetravalent and bivalent F<sub>v</sub> antibody constructs were determined from the A<sub>280</sub> values using the extinction coefficients  $\epsilon^{1\text{mg/ml}} = 1.96$  and 1.93, respectively.

**Example 4: Expression of the tetravalent and/or bivalent antibody construct in the yeast *Pichia pastoris***

Competent *P. pastoris* GS155 cells (Invitrogen) were electroporated in the presence of 10 µg plasmid-DNA of pPIC-DISC-LL and pPIC-DISC-SL, respectively, which had been linearized with SacI. The transformants were selected for 3 days at 30°C on YPD plates containing 100 µg/ml Zeocin™. The clones which secreted the bivalent and/or tetravalent F<sub>v</sub> antibody constructs were selected by plate screening using an anti-c-myc-mAb 9E10 (IC Chemikalien, Ismaning, Germany).

For the expression of the bivalent F<sub>v</sub> antibody constructs and tetravalent F<sub>v</sub> antibody constructs, respectively, the clones were cultured in YPD medium in shaking flasks for 2 days at 30°C with stirring. The cells were centrifuged resuspended in the same volume of the medium containing methanol and incubated for another 3 days at 30°C with stirring. The supernatants were obtained after the centrifugation. The recombinant product was isolated by ammonium sulfate precipitation, followed by IMAC as described above.

**Example 5: Characterization of the tetravalent F<sub>v</sub> antibody construct and bivalent F<sub>v</sub> antibody construct, respectively,**

(A) Size exclusion chromatography

An analytical gel filtration of the F<sub>v</sub> antibody constructs was carried out in PBS using a superdex 200-HR10/30 column (Pharmacia). The sample volume and the flow rate were 200 µl/min and 0.5 ml/min, respectively. The column was calibrated with high-molecular and low-molecular gel filtration calibration kits (Pharmacia).

#### (B) Flow cytometry

The human CD3<sup>+</sup>/CD19<sup>-</sup>-acute T-cell leukemia line Jurkat and the CD19<sup>+</sup>/CD3<sup>-</sup> B-cell line JOK-1 were used for flow cytometrie.  $5 \times 10^5$  cells in 50 µl RPMI 1640 medium (GIBCO BRL, Eggestein, Germany) which was supplemented with 10 % FCS and 0.1 % sodium azide (referred to as complete medium) were incubated with 100 µl of the F<sub>v</sub> antibody preparations for 45 minutes on ice. After washing using the complete medium the cells were incubated with 100 µl 10 µg/ml anti-c-myc-Mak 9E10 (IC Chemikalien) in the same buffer for 45 min on ice. After a second wash cycle, the cells were incubated with 100 µl of the FITC-labeled goat-anti-mouse-IgG (GIBCO BRL) under the same conditions as before. The cells were then washed again and resuspended in 100 µl 1 µg/ml propidium iodide solution (Sigma, Deisenhofen, Germany) in complete medium with the exclusion of dead cells. The relative fluorescence of the stained cells was measured using a FACScan flow cytometer (Becton Dickinson, Mountain View, CA).

#### (C) Cytotoxicity test

The CD19-expressing Burkitt lymphoma cell line Raji and Namalwa were used as target cells. The cells were incubated in RPMI 1640 (GIBCO BRL) which was supplemented with 10 %

heat-inactivated FCS (GIBCO BRL), 2 mM glutamine and 1 mM pyruvate, at 37°C in a dampened atmosphere with 7.5 % CO<sub>2</sub>. The cytotoxic T-cell tests were carried out in RPMI-1640 medium supplemented with 10 % FCS, 10 mM HEPES, 2 mM glutamine, 1 mM pyruvate and 0.05 mM 2-ME. The cytotoxic activity was evaluated using a standard [<sup>51</sup>Cr] release test; 2 x 10<sup>6</sup> target cells were labeled with 200 µCi Na[<sup>51</sup>Cr]O<sub>4</sub> (Amersham-Buchler, Braunschweig, Germany) and washed 4 times and then resuspended in medium in a concentration of 2 x 10<sup>5</sup>/ml. The effector cells were adjusted to a concentration of 5 x 10<sup>6</sup>/ml. Increasing amounts of CTLs in 100 µl were titrated to 10<sup>4</sup> target cells/well or cavity in 50 µl. 50 µl antibodies were added to each well. The entire test was prepared three times and incubated at 37°C for 4 h. 100 µl of the supernatant were collected and tested for [<sup>51</sup>Cr] release in a gamma counter (Cobra Auto Gamma; Canberra Packard, Dreieich, Germany). The maximum release was determined by incubation of the target cells in 10 % SDS, and the spontaneous release was determined by incubation of the cells in medium alone. The specific lysis (%) was calculated as: (experimental release - spontaneous release)/(maximum release - spontaneous release) x 100.

**Example 6: Construction of the plasmids pDISC5-LL and pDISC5-SL for the expression of bivalent, bispecific and/or tetravalent, bispecific F<sub>v</sub> antibody constructs in bacteria by high cell density fermentation**

Expression vectors were prepared which contained the hok/sok plasmid-free cell suicide system and a gene which codes for the Skp/OmpH periplasmic factor for a greater production of recombinant antibodies. The *skp* gene was amplified by PCR using the primers *skp*-1, 5'-CGA ATT CTT AAG ATA AGA AGG AGT

TTA TTG TGA AAA AGT GGT TAT TAG CTG CAG G and *skp-2*, 5'-CGA ATT AAG CTT CAT TAT TTA ACC TGT TTC AGT ACG TCG G using the plasmid pGAH317 (Holck and Kleppe, 1988, *Gene* 67, 117-124). The resulting PCR fragment was cleaved by AflIII and HindIII and inserted in the AflIII/HindIII-linearized plasmid pHKK (Horn et al., 1996, *Appl. Microbiol. Biotechnol.* 46, 524-532) so as to obtain the vector pSKK. The genes obtained in the plasmids pDISC3x19-LL and pDISC3x19-SL and coding for the scFv antibody constructs were amplified by means of the primers *fe-1*, 5'-CGA ATT TCT AGA TAA GAA GGA GAA ATT AAC CAT GAA ATA CC and *fe-2*, 5'-CGA ATT CTT AAG CTA TTA GTG ATG GTG ATG GTG ATG TGA G. The XbaI/AflIII-cleaved PCR fragments were inserted in pSKK before the *skp* insert so as to obtain the expression plasmids pDISC5-LL and pDISC6-SL, respectively, which contain tri-cistronic operons under the control of the *lac* promoter/operator system (cf. figs. 9, 10).

## SEQUENCE RECORD

- (1) GENERAL INDICATIONS:
- (i) APPLICANT:
    - (A) NAME: Deutsches Krebsforschungszentrum
    - (B) STREET: Im Neuenheimer Feld 280
    - (C) TOWN: Heidelberg
    - (E) COUNTRY: Germany
    - (F) POSTAL CODE: 69120
  - (ii) TITLE OF THE INVENTION: Multivalent Antibody Constructs
  - (iii) NUMBER OF SEQUENCES: 17
  - (iv) COMPUTER-READABLE VERSION:
    - (A) DATA CARRIER: floppy disk
    - (B) COMPUTER: IBM PC compatible
    - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
    - (D) SOFTWARE: PatentIn Release #1.0, version #1.30 (EPA)
- (2) INDICATIONS AS TO SEQ ID NO: 1:
- (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 1698 base pairs
    - (B) KIND: nucleotide
    - (C) STRAND TYPE: single strand
    - (D) TOPOLOGY: linear
  - (ii) KIND OF MOLECULE: genome DNA
  - (iii) HYPOTHETICAL: no
  - (iv) ANTISENSE: no
  - (ix) FEATURE:
    - (A) NAME/KEY: CDS
    - (B) POSITION: 28..1689
  - (ix) FEATURE:
    - (A) NAME/KEY: mat\_peptide
    - (B) POSITION: 28..1689
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GAATTCATTA AAGAGGAGAA ATTAACC ATG AAA TAC CTA TTG CCT ACG GCA  
 Met Lys Tyr Leu Leu Pro Thr Ala  
 1 5

2

GCC GCT GGC TTG CTG CTG CTG GCA GCT CAG CCG GCC ATG GCG CAG GTG Ala Ala Gly Leu Leu Leu Leu Ala Ala Gln Pro Ala Met Ala Gln Val 10 15 20	99
CAA CTG CAG CAG TCT GGG GCT GAA CTG GCA AGA CCT GGG GCC TCA GTG Gln Leu Gln Gln Ser Gly Ala Glu Leu Ala Arg Pro Gly Ala Ser Val 25 30 35 40	147
AAG ATG TCC TGC AAG GCT TCT GGC TAC ACC TTT ACT AGG TAC ACG ATG Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Arg Tyr Thr Met- 45 50 55	195
CAC TGG GTA AAA CAG AGG CCT GGA CAG GGT CTG GAA TGG ATT GGA TAC His Trp Val Lys Gln Arg Pro Gly Gln Gly Leu Glu Trp Ile Gly Tyr 60 65 70	243
ATT AAT CCT AGC CGT GGT TAT ACT AAT TAC AAT CAG AAG TTC AAG GAC Ile Asn Pro Ser Arg Gly Tyr Thr Asn Tyr Asn Gln Lys Phe Lys Asp 75 80 85	291
AAG GCC ACA TTG ACT ACA GAC AAA TCC TCC AGC ACA GCC TAC ATG CAA Lys Ala Thr Leu Thr Thr Asp Lys Ser Ser Ser Thr Ala Tyr Met Gln 90 95 100	339
CTG AGC AGC CTG ACA TCT GAG GAC TCT GCA GTC TAT TAC TGT GCA AGA Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Tyr Cys Ala Arg 105 110 115 120	387
TAT TAT GAT GAT CAT TAC AGC CTT GAC TAC TGG GGC CAA GGC ACC ACT Tyr Tyr Asp Asp His Tyr Ser Leu Asp Tyr Trp Gly Gln Gly Thr Thr 125 130 135	435
CTC ACA GTC TCC TCA GCC AAA ACA ACA CCC AAG CTT GGC GGT GAT ATC Leu Thr Val Ser Ser Ala Lys Thr Thr Pro Lys Leu Gly Gly Asp Ile 140 145 150	483
TTG CTC ACC CAA ACT CCA GCT TCT TTG GCT GTG TCT CTA GGG CAG AGG Leu Leu Thr Gln Thr Pro Ala Ser Leu Ala Val Ser Leu Gly Gln Arg 155 160 165	531
GCC ACC ATC TCC TGC AAG GCC AGC CAA AGT GTT GAT TAT GAT GGT GAT Ala Thr Ile Ser Cys Lys Ala Ser Gln Ser Val Asp Tyr Asp Gly Asp 170 175 180	579
AGT TAT TTG AAC TGG TAC CAA CAG ATT CCA GGA CAG CCA CCC AAA CTC Ser Tyr Leu Asn Trp Tyr Gln Gln Ile Pro Gly Gln Pro Pro Lys Leu 185 190 195 200	627
CTC ATC TAT GAT GCA TCC AAT CTA GTT TCT GGG ATC CCA CCC AGG TTT Leu Ile Tyr Asp Ala Ser Asn Leu Val Ser Gly Ile Pro Pro Arg Phe 205 210 215	675
AGT GGC AGT GGG TCT GGG ACA GAC TTC ACC CTC AAC ATC CAT CCT GTG Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Asn Ile His Pro Val 220 225 230	723

GAG AAG GTG GAT GCT GCA ACC TAT CAC TGT CAG CAA AGT ACT GAG GAT Glu Lys Val Asp Ala Ala Thr Tyr His Cys Gln Gln Ser Thr Glu Asp 235 240 245	771
CCG TGG ACG TTC GGT GGA GGC ACC AAG CTG GAA ATC AAA CGG GCT GAT Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg Ala Asp 250 255 260	819
GCT GCG GCC GCT GGT GGT GGT GGT TCT GGC GGC GGT GGT AGC GGT GGT Ala Ala Ala Ala Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly 265 270 275 280	867
GGC GGC TCC GGT GGT GGT GGT AGC CAG GTG CAG CTG CAG CAG TCT GGG Gly Gly Ser Gly Gly Gly Gly Ser Gln Val Gln Leu Gln Gln Ser Gly 285 290 295	915
GCT GAG CTG GTG AGG CCT GGG TCC TCA GTG AAG ATT TCC TGC AAG GCT Ala Glu Leu Val Arg Pro Gly Ser Ser Val Lys Ile Ser Cys Lys Ala 300 305 310	963
TCT GGC TAT GCA TTC AGT AGC TAC TGG ATG AAC TGG GTG AAG CAG AGG Ser Gly Tyr Ala Phe Ser Ser Tyr Trp Met Asn Trp Val Lys Gln Arg 315 320 325	1011
CCT GGA CAG GGT CTT GAG TGG ATT GGA CAG ATT TGG CCT GGA GAT GGT Pro Gly Gln Gly Leu Glu Trp Ile Gly Gln Ile Trp Pro Gly Asp Gly 330 335 340	1059
GAT ACT AAC TAC AAT GGA AAG TTC AAG GGT AAA GCC ACT CTG ACT GCA Asp Thr Asn Tyr Asn Gly Lys Phe Lys Gly Lys Ala Thr Leu Thr Ala 345 350 355 360	1107
GAC GAA TCC TCC AGC ACA GCC TAC ATG CAA CTC AGC AGC CTA GCA TCT Asp Glu Ser Ser Ser Thr Ala Tyr Met Gln Leu Ser Ser Leu Ala Ser 365 370 375	1155
GAG GAC TCT GCG GTC TAT TTC TGT GCA AGA CGG GAG ACT ACG ACG GTA Glu Asp Ser Ala Val Tyr Phe Cys Ala Arg Arg Glu Thr Thr Thr Val 380 385 390	1203
GGC CGT TAT TAC TAT GCT ATG GAC TAC TGG GGT CAA GGA ACC TCA GTC Gly Arg Tyr Tyr Tyr Ala Met Asp Tyr Trp Gly Gln Gly Thr Ser Val 395 400 405	1251
ACC GTC TCC TCA GCC AAA ACA ACA CCC AAG CTT GGC GGT GAT ATC GTG Thr Val Ser Ser Ala Lys Thr Thr Pro Lys Leu Gly Gly Asp Ile Val 410 415 420	1299
CTC ACT CAG TCT CCA GCA ATC ATG TCT GCA TCT CCA GGG GAG AAG GTC Leu Thr Gln Ser Pro Ala Ile Met Ser Ala Ser Pro Gly Glu Lys Val 425 430 435 440	1347
ACC ATG ACC TGC AGT GCC AGC TCA AGT GTA AGT TAC ATG AAC TGG TAC Thr Met Thr Cys Ser Ala Ser Ser Ser Val Ser Tyr Met Asn Trp Tyr 445 450 455	1395

CAG CAG AAG TCA GGC ACC TCC CCC AAA AGA TGG ATT TAT GAC ACA TCC Gln Gln Lys Ser Gly Thr Ser Pro Lys Arg Trp Ile Tyr Asp Thr Ser 460 465 470	1443
AAA CTG GCT TCT GGA GTC CCT GCT CAC TTC AGG GGC AGT GGG TCT GGG Lys Leu Ala Ser Gly Val Pro Ala His Phe Arg Gly Ser Gly Ser Gly 475 480 485	1491
ACC TCT TAC TCT CTC ACA ATC AGC GGC ATG GAG GCT GAA GAT GCT GCC Thr Ser Tyr Ser Leu Thr Ile Ser Gly Met Glu Ala Glu Asp Ala Ala 490 495 500	1539
ACT TAT TAC TGC CAG CAG TGG AGT AGT AAC CCA TTC ACG TTC GGC TCG Thr Tyr Tyr Cys Gln Gln Trp Ser Ser Asn Pro Phe Thr Phe Gly Ser 505 510 515 520	1587
GGC ACA AAG TTG GAA ATA AAC CGG GCT GAT ACT GCA CCA ACT GGA TCC Gly Thr Lys Leu Glu Ile Asn Arg Ala Asp Thr Ala Pro Thr Gly Ser 525 530 535	1635
GAA CAA AAG CTG ATC TCA GAA GAA GAC CTA AAC TCA CAT CAC CAT CAC Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn Ser His His His His 540 545 550	1683
CAT CAC TAATCTAGA His His	1698

## (2) INDICATIONS AS TO ID NO: 2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 554 amino acids
- (B) KIND: amino acid
- (D) TOPOLOGY: linear

## (ii) KIND OF MOLECULE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala 1 5 10 15
Ala Gln Pro Ala Met Ala Gln Val Gln Leu Gln Gln Ser Gly Ala Glu 20 25 30
Leu Ala Arg Pro Gly Ala Ser Val Lys Met Ser Cys Lys Ala Ser Gly 35 40 45
Tyr Thr Phe Thr Arg Tyr Thr Met His Trp Val Lys Gln Arg Pro Gly 50 55 60
Gln Gly Leu Glu Trp Ile Gly Tyr Ile Asn Pro Ser Arg Gly Tyr Thr 65 70 75 80



Asn Tyr Asn Gln Lys Phe Lys Asp Lys Ala Thr Leu Thr Thr Asp Lys  
                     85                    90                    95  
 Ser Ser Ser Thr Ala Tyr Met Gln Leu Ser Ser Leu Thr Ser Glu Asp  
                     100                    105                    110  
 Ser Ala Val Tyr Tyr Cys Ala Arg Tyr Tyr Asp Asp His Tyr Ser Leu  
                     115                    120                    125  
 Asp Tyr Trp Gly Gln Gly Thr Thr Leu Thr Val Ser Ser Ala Lys Thr  
                     130                    135                    140  
 Thr Pro Lys Leu Gly Gly Asp Ile Leu Leu Thr Gln Thr Pro Ala Ser  
                     145                    150                    155                    160  
 Leu Ala Val Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys Lys Ala Ser  
                     165                    170                    175  
 Gln Ser Val Asp Tyr Asp Gly Asp Ser Tyr Leu Asn Trp Tyr Gln Gln  
                     180                    185                    190  
 Ile Pro Gly Gln Pro Pro Lys Leu Leu Ile Tyr Asp Ala Ser Asn Leu  
                     195                    200                    205  
 Val Ser Gly Ile Pro Pro Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp  
                     210                    215                    220  
 Phe Thr Leu Asn Ile His Pro Val Glu Lys Val Asp Ala Ala Thr Tyr  
                     225                    230                    235                    240  
 His Cys Gln Gln Ser Thr Glu Asp Pro Trp Thr Phe Gly Gly Gly Thr  
                     245                    250                    255  
 Lys Leu Glu Ile Lys Arg Ala Asp Ala Ala Ala Gly Gly Gly Gly  
                     260                    265                    270  
 Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser  
                     275                    280                    285  
 Gln Val Gln Leu Gln Gln Ser Gly Ala Glu Leu Val Arg Pro Gly Ser  
                     290                    295                    300  
 Ser Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Ala Phe Ser Ser Tyr  
                     305                    310                    315                    320  
 Trp Met Asn Trp Val Lys Gln Arg Pro Gly Gln Gly Leu Glu Trp Ile  
                     325                    330                    335  
 Gly Gln Ile Trp Pro Gly Asp Gly Asp Thr Asn Tyr Asn Gly Lys Phe  
                     340                    345                    350  
 Lys Gly Lys Ala Thr Leu Thr Ala Asp Glu Ser Ser Ser Thr Ala Tyr  
                     355                    360                    365

6

Met Gln Leu Ser Ser Leu Ala Ser Glu Asp Ser Ala Val Tyr Phe Cys  
 370 375 380  
 Ala Arg Arg Glu Thr Thr Thr Val Gly Arg Tyr Tyr Tyr Ala Met Asp  
 385 390 395 400  
 Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser Ala Lys Thr Thr  
 405 410 415  
 Pro Lys Leu Gly Gly Asp Ile Val Leu Thr Gln Ser Pro Ala Ile Met  
 420 425 430  
 Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Ser Ala Ser Ser  
 435 440 445  
 Ser Val Ser Tyr Met Asn Trp Tyr Gln Gln Lys Ser Gly Thr Ser Pro  
 450 455 460  
 Lys Arg Trp Ile Tyr Asp Thr Ser Lys Leu Ala Ser Gly Val Pro Ala  
 465 470 475 480  
 His Phe Arg Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile Ser  
 485 490 495  
 Gly Met Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp Ser  
 500 505 510  
 Ser Asn Pro Phe Thr Phe Gly Ser Gly Thr Lys Leu Glu Ile Asn Arg  
 515 520 525  
 Ala Asp Thr Ala Pro Thr Gly Ser Glu Gln Lys Leu Ile Ser Glu Glu  
 530 535 540  
 Asp Leu Asn Ser His His His His His  
 545 550

- (2) INDICATIONS AS TO ID NO: 3:
- (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 1653 base pairs
    - (B) KIND: nucleotide
    - (C) STRAND TYPE: single strand
    - (D) TOPOLOGY: linear
  - (ii) KIND OF MOLECULE: genome DNA
  - (iii) HYPOTHETICAL: no
  - (iv) ANTISENSE: no
  - (ix) FEATURE:
    - (A) NAME/KEY: CDS
    - (B) POSITION: 28..1644

## (ix) FEATURE:

(A) NAME/KEY: mat\_peptide

(B) POSITION: 28..1644

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GAATTCATTA AAGAGGAGAA ATTAACC ATG AAA TAC CTA TTG CCT ACG GCA	51
Met Lys Tyr Leu Leu Pro Thr Ala	
1 5	
GCC GCT GGC TTG CTG CTG CTG GCA GCT CAG CCG GCC ATG GCG CAG GTG	99
Ala Ala Gly Leu Leu Leu Ala Ala Gln Pro Ala Met Ala Gln Val	
10 15 20	
CAA CTG CAG CAG TCT GGG GCT GAA CTG GCA AGA CCT GGG GCC TCA GTG	147
Gln Leu Gln Gln Ser Gly Ala Glu Leu Ala Arg Pro Gly Ala Ser Val	
25 30 35 40	
AAG ATG TCC TGC AAG GCT TCT GGC TAC ACC TTT ACT AGG TAC ACG ATG	195
Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Arg Tyr Thr Met	
45 50 55	
CAC TGG GTA AAA CAG AGG CCT GGA CAG GGT CTG GAA TGG ATT GGA TAC	243
His Trp Val Lys Gln Arg Pro Gly Gln Gly Leu Glu Trp Ile Gly Tyr	
60 65 70	
ATT AAT CCT AGC CGT GGT TAT ACT AAT TAC AAT CAG AAG TTC AAG GAC	291
Ile Asn Pro Ser Arg Gly Tyr Thr Asn Tyr Asn Gln Lys Phe Lys Asp	
75 80 85	
AAG GCC ACA TTG ACT ACA GAC AAA TCC TCC AGC ACA GCC TAC ATG CAA	339
Lys Ala Thr Leu Thr Thr Asp Lys Ser Ser Ser Thr Ala Tyr Met Gln	
90 95 100	
CTG AGC AGC CTG ACA TCT GAG GAC TCT GCA GTC TAT TAC TGT GCA AGA	387
Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Tyr Cys Ala Arg	
105 110 115 120	
TAT TAT GAT GAT CAT TAC AGC CTT GAC TAC TGG GGC CAA GGC ACC ACT	435
Tyr Tyr Asp Asp His Tyr Ser Leu Asp Tyr Trp Gly Gln Gly Thr Thr	
125 130 135	
CTC ACA GTC TCC TCA GCC AAA ACA ACA CCC AAG CTT GGC GGT GAT ATC	483
Leu Thr Val Ser Ser Ala Lys Thr Thr Pro Lys Leu Gly Gly Asp Ile	
140 145 150	
TTG CTC ACC CAA ACT CCA GCT TCT TTG GCT GTG TCT CTA GGG CAG AGG	531
Leu Leu Thr Gln Thr Pro Ala Ser Leu Ala Val Ser Leu Gly Gln Arg	
155 160 165	
GCC ACC ATC TCC TGC AAG GCC AGC CAA AGT GTT GAT TAT GAT GGT GAT	579
Ala Thr Ile Ser Cys Lys Ala Ser Gln Ser Val Asp Tyr Asp Gly Asp	
170 175 180	

AGT TAT TTG AAC TGG TAC CAA CAG ATT CCA GGA CAG CCA CCC AAA CTC Ser Tyr Leu Asn Trp Tyr Gln Gln Ile Pro Gly Gln Pro Pro Lys Leu 185 190 195 200	627
CTC ATC TAT GAT GCA TCC AAT CTA GTT TCT GGG ATC CCA CCC AGG TTT Leu Ile Tyr Asp Ala Ser Asn Leu Val Ser Gly Ile Pro Pro Arg Phe 205 210 215	675
AGT GGC AGT GGG TCT GGG ACA GAC TTC ACC CTC AAC ATC CAT CCT GTG Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Asn Ile His Pro Val 220 225 230	723
GAG AAG GTG GAT GCT GCA ACC TAT CAC TGT CAG CAA AGT ACT GAG GAT Glu Lys Val Asp Ala Ala Thr Tyr His Cys Gln Gln Ser Thr Glu Asp 235 240 245	771
CCG TGG ACG TTC GGT GGA GGC ACC AAG CTG GAA ATC AAA CGG GCT GAT Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg Ala Asp 250 255 260	819
GCT GCG GCC GCT GGT GGC CCA GGG TCG CAG GTG CAG CTG CAG CAG TCT Ala Ala Ala Ala Gly Gly Pro Gly Ser Gln Val Gln Leu Gln Gln Ser 265 270 275 280	867
GGG GCT GAG CTG GTG AGG CCT GGG TCC TCA GTG AAG ATT TCC TGC AAG Gly Ala Glu Leu Val Arg Pro Gly Ser Ser Val Lys Ile Ser Cys Lys 285 290 295	915
GCT TCT GGC TAT GCA TTC AGT AGC TAC TGG ATG AAC TGG GTG AAG CAG Ala Ser Gly Tyr Ala Phe Ser Ser Tyr Trp Met Asn Trp Val Lys Gln 300 305 310	963
AGG CCT GGA CAG GGT CTT GAG TGG ATT GGA CAG ATT TGG CCT GGA GAT Arg Pro Gly Gln Gly Leu Glu Trp Ile Gly Gln Ile Trp Pro Gly Asp 315 320 325	1011
GGT GAT ACT AAC TAC AAT GGA AAG TTC AAG GGT AAA GCC ACT CTG ACT Gly Asp Thr Asn Tyr Asn Gly Lys Phe Lys Gly Lys Ala Thr Leu Thr 330 335 340	1059
GCA GAC GAA TCC TCC AGC ACA GCC TAC ATG CAA CTC AGC AGC CTA GCA Ala Asp Glu Ser Ser Ser Thr Ala Tyr Met Gln Leu Ser Ser Leu Ala 345 350 355 360	1107
TCT GAG GAC TCT GCG GTC TAT TTC TGT GCA AGA CGG GAG ACT ACG ACG Ser Glu Asp Ser Ala Val Tyr Phe Cys Ala Arg Arg Glu Thr Thr Thr 365 370 375	1155
GTA GGC CGT TAT TAC TAT GCT ATG GAC TAC TGG GGT CAA GGA ACC TCA Val Gly Arg Tyr Tyr Tyr Ala Met Asp Tyr Trp Gly Gln Gly Thr Ser 380 385 390	1203
GTC ACC GTC TCC TCA GCC AAA ACA ACA CCC AAG CTT GGC GGT GAT ATC Val Thr Val Ser Ser Ala Lys Thr Thr Pro Lys Leu Gly Gly Asp Ile 395 400 405	1251

9

GTG CTC ACT CAG TCT CCA GCA ATC ATG TCT GCA TCT CCA GGG GAG AAG	1299
Val Leu Thr Gln Ser Pro Ala Ile Met Ser Ala Ser Pro Gly Glu Lys	
410 415 420	
GTC ACC ATG ACC TGC AGT GCC AGC TCA AGT GTA AGT TAC ATG AAC TGG	1347
Val Thr Met Thr Cys Ser Ala Ser Ser Ser Val Ser Tyr Met Asn Trp	
425 430 435 440	
TAC CAG CAG AAG TCA GGC ACC TCC CCC AAA AGA TGG ATT TAT GAC ACA	1395
Tyr Gln Gln Lys Ser Gly Thr Ser Pro Lys Arg Trp Ile Tyr Asp Thr	
445 450 455	
TCC AAA CTG GCT TCT GGA GTC CCT GCT CAC TTC AGG GGC AGT GGG TCT	1443
Ser Lys Leu Ala Ser Gly Val Pro Ala His Phe Arg Gly Ser Gly Ser	
460 465 470	
GGG ACC TCT TAC TCT CTC ACA ATC AGC GGC ATG GAG GCT GAA GAT GCT	1491
Gly Thr Ser Tyr Ser Leu Thr Ile Ser Gly Met Glu Ala Glu Asp Ala	
475 480 485	
GCC ACT TAT TAC TGC CAG CAG TGG AGT AGT AAC CCA TTC ACG TTC GGC	1539
Ala Thr Tyr Tyr Cys Gln Gln Trp Ser Ser Asn Pro Phe Thr Phe Gly	
490 495 500	
TCG GGG ACA AAG TTG GAA ATA AAC CGG GCT GAT ACT GCA CCA ACT GGA	1587
Ser Gly Thr Lys Leu Glu Ile Asn Arg Ala Asp Thr Ala Pro Thr Gly	
505 510 515 520	
TCC GAA CAA AAG CTG ATC TCA GAA GAA GAC CTA AAC TCA CAT CAC CAT	1635
Ser Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn Ser His His His	
525 530 535	
CAC CAT CAC TAATCTAGA	1653
His His His	

## (2) INDICATIONS AS TO ID NO: 4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 539 amino acids
- (B) KIND: amino acid
- (D) TOPOLOGY: linear

## (ii) KIND OF MOLECULE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
1 5 10 15
Ala Gln Pro Ala Met Ala Gln Val Gln Leu Gln Gln Ser Gly Ala Glu
20 25 30
Leu Ala Arg Pro Gly Ala Ser Val Lys Met Ser Cys Lys Ala Ser Gly
35 40 45

10

Tyr Thr Phe Thr Arg Tyr Thr Met His Trp Val Lys Gln Arg Pro Gly  
 50 55 60  
 Gln Gly Leu Glu Trp Ile Gly Tyr Ile Asn Pro Ser Arg Gly Tyr Thr  
 65 70 75 80  
 Asn Tyr Asn Gln Lys Phe Lys Asp Lys Ala Thr Leu Thr Thr Asp Lys  
 85 90 95  
 Ser Ser Ser Thr Ala Tyr Met Gln Leu Ser Ser Leu Thr Ser Glu Asp  
 100 105 110  
 Ser Ala Val Tyr Tyr Cys Ala Arg Tyr Tyr Asp Asp His Tyr Ser Leu  
 115 120 125  
 Asp Tyr Trp Gly Gln Gly Thr Thr Leu Thr Val Ser Ser Ala Lys Thr  
 130 135 140  
 Thr Pro Lys Leu Gly Gly Asp Ile Leu Leu Thr Gln Thr Pro Ala Ser  
 145 150 155 160  
 Leu Ala Val Ser Leu Gly Gln Arg Ala Thr Ile Ser Cys Lys Ala Ser  
 165 170 175  
 Gln Ser Val Asp Tyr Asp Gly Asp Ser Tyr Leu Asn Trp Tyr Gln Gln  
 180 185 190  
 Ile Pro Gly Gln Pro Pro Lys Leu Leu Ile Tyr Asp Ala Ser Asn Leu  
 195 200 205  
 Val Ser Gly Ile Pro Pro Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp  
 210 215 220  
 Phe Thr Leu Asn Ile His Pro Val Glu Lys Val Asp Ala Ala Thr Tyr  
 225 230 235 240  
 His Cys Gln Gln Ser Thr Glu Asp Pro Trp Thr Phe Gly Gly Gly Thr  
 245 250 255  
 Lys Leu Glu Ile Lys Arg Ala Asp Ala Ala Ala Ala Gly Gly Pro Gly  
 260 265 270  
 Ser Gln Val Gln Leu Gln Gln Ser Gly Ala Glu Leu Val Arg Pro Gly  
 275 280 285  
 Ser Ser Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Ala Phe Ser Ser  
 290 295 300  
 Tyr Trp Met Asn Trp Val Lys Gln Arg Pro Gly Gln Gly Leu Glu Trp  
 305 310 315 320  
 Ile Gly Gln Ile Trp Pro Gly Asp Gly Asp Thr Asn Tyr Asn Gly Lys  
 325 330 335

Phe Lys Gly Lys Ala Thr Leu Thr Ala Asp Glu Ser Ser Ser Thr Ala  
 340 345 350  
 Tyr Met Gln Leu Ser Ser Leu Ala Ser Glu Asp Ser Ala Val Tyr Phe  
 355 360 365  
 Cys Ala Arg Arg Glu Thr Thr Thr Val Gly Arg Tyr Tyr Tyr Ala Met  
 370 375 380  
 Asp Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser Ala Lys Thr  
 385 390 395 400  
 Thr Pro Lys Leu Gly Gly Asp Ile Val Leu Thr Gln Ser Pro Ala Ile  
 405 410 415  
 Met Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Ser Ala Ser  
 420 425 430  
 Ser Ser Val Ser Tyr Met Asn Trp Tyr Gln Gln Lys Ser Gly Thr Ser  
 435 440 445  
 Pro Lys Arg Trp Ile Tyr Asp Thr Ser Lys Leu Ala Ser Gly Val Pro  
 450 455 460  
 Ala His Phe Arg Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile  
 465 470 475 480  
 Ser Gly Met Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp  
 485 490 495  
 Ser Ser Asn Pro Phe Thr Phe Gly Ser Gly Thr Lys Leu Glu Ile Asn  
 500 505 510  
 Arg Ala Asp Thr Ala Pro Thr Gly Ser Glu Gln Lys Leu Ile Ser Glu  
 515 520 525  
 Glu Asp Leu Asn Ser His His His His His  
 530 535

- (2) INDICATIONS AS TO ID NO: 5:
- (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 57 base pairs
    - (B) KIND: nucleotide
    - (C) STRAND TYPE: single strand
    - (D) TOPOLOGY: linear
  - (ii) KIND OF MOLECULE: other nucleic acid
    - (A) DESCRIPTION: /desc = "primer"
  - (iii) HYPOTHETICAL: no
  - (iv) ANTISENSE: no
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

TATATACTGC AGCTGCACCT GCGACCTTGG GCCACCAGCG GCCGCAGCAT CAGCCCCG

57

(2) INDICATIONS AS TO ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 45 base pairs
  - (B) KIND: nucleotide
  - (C) STRAND TYPE: single strand
  - (D) TOPOLOGY: linear

(ii) KIND OF MOLECULE: other nucleic acid

(A) DESCRIPTION: /desc = "primer"

(iii) HYPOTHETICAL: no

(iv) ANTISENSE: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

CCGTGAATTC CAGGTGCAAC TGCAGCAGTC TGGGGCTGAA CTGGC

45

(2) INDICATIONS AS TO ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 34 base pairs
  - (B) KIND: nucleotide
  - (C) STRAND TYPE: single strand
  - (D) TOPOLOGY: linear
- (ii) KIND OF MOLECULE: other nucleic acid
  - (A) DESCRIPTION: /desc = "primer"
- (iii) HYPOTHETICAL: no
- (iv) ANTISENSE: no
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GGTCGACGTT AACCGACAAA CAACAGATAA AACG

34



## (2) INDICATIONS AS TO ID NO: 8:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 348 base pairs
- (B) KIND: nucleotide
- (C) STRAND TYPE: single strand
- (D) TOPOLOGY: linear

(ii) KIND OF MOLECULE: genome DNA

(iii) HYPOTHETICAL: no

(iv) ANTISENSE: no

## (ix) FEATURE:

(A) NAME/KEY: CDS

(B) POSITION: 1..348

## (ix) FEATURE:

(A) NAME/KEY: mat\_peptide

(B) POSITION: 1..348

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

ATG AGA TTT CCT TCA ATT TTT ACT GCT GTT TTA TTC GCA GCA TCC TCC	48
Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser	
1 5 10 15	
GCA TTA GCT GCT CCA GTC AAC ACT ACA ACA GAA GAT GAA ACG GCA CAA	96
Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala Gln	
20 25 30	
ATT CCG GCT GAA GCT GTC ATC GGT TAC TCA GAT TTA GAA GGG GAT TTC	144
Ile Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe	
35 40 45	
GAT GTT GCT GTT TTG CCA TTT TCC AAC AGC ACA AAT AAC GGG TTA TTG	192
Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu	
50 55 60	
TTT ATA AAT ACT ACT ATT GCC AGC ATT GCT GCT AAA GAA GAA GGG GTA	240
Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val	
65 70 75 80	
TCT CTC GAG AAA AGA GAG GCT GAA GCT GAA TTC CAG GTG CAA CTG CAG	288
Ser Leu Glu Lys Arg Glu Ala Glu Ala Glu Phe Gln Val Gln Leu Gln	
85 90 95	
CAG TCT GGG GCT GAA CTG GCA AGA CCT GGG GCC TCA GTG AAG ATG TCC	336
Gln Ser Gly Ala Glu Leu Ala Arg Pro Gly Ala Ser Val Lys Met Ser	
100 105 110	
TGC AAG GCT TCT	348
Cys Lys Ala Ser	
115	

14

## 2) INDICATIONS AS TO ID NO: 9:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 116 amino acids
- (B) KIND: amino acid
- (D) TOPOLOGY: linear

## (ii) KIND OF MOLECULE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

```

Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser
 1           5           10           15
Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala Gln
          20           25           30
Ile Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe
          35           40           45
Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu
          50           55           60
Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val
          65           70           75           80
Ser Leu Glu Lys Arg Glu Ala Glu Ala Glu Phe Gln Val Gln Leu Gln
          85           90           95
Gln Ser Gly Ala Glu Leu Ala Arg Pro Gly Ala Ser Val Lys Met Ser
          100          105          110
Cys Lys Ala Ser
          115

```

## (2) INDICATIONS AS TO ID NO: 10:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 354 base pairs
- (B) KIND: nucleotide
- (C) STRAND TYPE: single strand
- (D) TOPOLOGY: linear

## (ii) KIND OF MOLECULE: genome DNA

## (iii) HYPOTHETICAL: no

## (iv) ANTISENSE: no

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) POSITION: 1..354

## (ix) FEATURE:

- (A) NAME/KEY: mat\_peptide
- (B) POSITION: 1..354

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

15

ATG AGA TTT CCT TCA ATT TTT ACT GCT GTT TTA TTC GCA GCA TCC TCC	48
Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser	
1 5 10 15	
GCA TTA GCT GCT CCA GTC AAC ACT ACA ACA GAA GAT GAA ACG GCA CAA	96
Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala Gln	
20 25 30	
ATT CCG GCT GAA GCT GTC ATC GGT TAC TCA GAT TTA GAA GGG GAT TTC	144
Ile Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe	
35 40 45	
GAT GTT GCT GTT TTG CCA TTT TCC AAC AGC ACA AAT AAC GGG TTA TTG	192
Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu	
50 55 60	
TTT ATA AAT ACT ACT ATT GCC AGC ATT GCT GCT AAA GAA GAA GGG GTA	240
Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val	
65 70 75 80	
TCT CTC GAG AAA AGA GAG GCT GAA GCT GAA TTC ATG GCG CAG GTG CAA	288
Ser Leu Glu Lys Arg Glu Ala Glu Ala Glu Phe Met Ala Gln Val Gln	
85 90 95	
CTG CAG CAG TCT GGG GCT GAA CTG GCA AGA CCT GGG GCC TCA GTG AAG	336
Leu Gln Gln Ser Gly Ala Glu Leu Ala Arg Pro Gly Ala Ser Val Lys	
100 105 110	
ATG TCC TGC AAG GCT TCT	354
Met Ser Cys Lys Ala Ser	
115	

## 2) INDICATIONS AS TO ID NO:11:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 118 amino acids
- (B) KIND: amino acid
- (D) TOPOLOGY: linear

## (ii) KIND OF MOLECULE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser
1 5 10 15
Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala Gln
20 25 30
Ile Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe
35 40 45

16

```

Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu
  50                               55                               60

Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val
  65                               70                               75                               80

Ser Leu Glu Lys Arg Glu Ala Glu Ala Glu Phe Met Ala Gln Val Gln
                        85                               90                               95

Leu Gln Gln Ser Gly Ala Glu Leu Ala Arg Pro Gly Ala Ser Val Lys
      100                               105                               110

Met Ser Cys Lys Ala Ser
      115

```

- (2) INDICATIONS AS TO ID NO: 12:
- (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 42 base pairs
    - (B) KIND: nucleotide
    - (C) STRAND TYPE: single strand
    - (D) TOPOLOGY: linear
  - (ii) KIND OF MOLECULE: other nucleic acid
    - (A) DESCRIPTION: /desc = "primer"
  - (iii) HYPOTHETICAL: no
  - (iv) ANTISENSE: no
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

TCACACAGAA TTCTTAGATC TATTAAAGAG GAGAAATTAA CC

42

- (2) INDICATIONS AS TO ID NO: 13:
- (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 40 base pairs
    - (B) KIND: nucleotide
    - (C) STRAND TYPE: single strand
    - (D) TOPOLOGY: linear
  - (ii) KIND OF MOLECULE: other nucleic acid
    - (A) DESCRIPTION: /desc = "primer"
  - (iii) HYPOTHETICAL: no
  - (iv) ANTISENSE: no
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

AGCACACGAT ATCACCGCCA AGCTTGGGTG TTGTTTTGGC

40

- (2) INDICATIONS AS TO ID NO: 14:
- (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 43 base pairs
    - (B) KIND: nucleotide
    - (C) STRAND TYPE: single strand
    - (D) TOPOLOGY: linear
  - (ii) KIND OF MOLECULE: other nucleic acid
    - (A) DESCRIPTION: /desc = "primer"
  - (iii) HYPOTHETICAL: no
  - (iv) ANTISENSE: no
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14

AGCACACAAG CTTGGCGGTG ATATCTTGCT CACCCAACT CCA

43

- (2) INDICATIONS AS TO ID NO: 15:
- (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 57 base pairs
    - (B) KIND: nucleotide
    - (C) STRAND TYPE: single strand
    - (D) TOPOLOGY: linear
  - (ii) KIND OF MOLECULE: other nucleic acid
    - (A) DESCRIPTION: /desc = "primer"
  - (iii) HYPOTHETICAL: no
  - (iv) ANTISENSE: no
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15

AGCACACTCT AGAGACACAC AGATCTTTAG TGATGGTGAT GGTGATGTGA GTTTAGG

57

- (2) INDICATIONS AS TO ID NO: 16:
- (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 33 base pairs
    - (B) KIND: nucleotide
    - (C) STRAND TYPE: single strand
    - (D) TOPOLOGY: linear

18

- (ii) KIND OF MOLECULE: other nucleic acid
  - (A) DESCRIPTION: /desc = "primer"
- (iii) HYPOTHETICAL: no
- (iv) ANTISENSE: no
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

CAGCCGGCCA TGGCGCAGGT GCAACTGCAG CAG

33

- (2) INDICATIONS AS TO ID NO: 17:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 102 base pairs
    - (B) KIND: nucleotide
    - (C) STRAND TYPE: single strand
    - (D) TOPOLOGY: linear
  - (ii) KIND OF MOLECULE: other nucleic acid
    - (A) DESCRIPTION: /desc = "primer"
  - (iii) HYPOTHETICAL: no
  - (iv) ANTISENSE: no
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

TATATACTGC AGCTGCACCT GGCTACCACC ACCACCGGAG CCGCCACCAC CGCTACCACC

60

GCCGCCAGAA CCACCACCAC CAGCGGCCGC AGCATCAGCC CG

102

Official File: PCT/DE99/01350

Attorney's File: K 2675

### **Amended Claims**

1. A multivalent F<sub>v</sub> antibody construct having at least four variable domains which are linked with one another via the peptide linkers 1, 2 and 3, wherein the peptide linkers 1 and 3 have 0 to 10 amino acids.
2. The F<sub>v</sub> antibody construct according to claim 1, wherein the peptide linkers 1 and 3 have the amino acid sequence GG.
3. The F<sub>v</sub> antibody construct according to claim 1 or 2, wherein the F<sub>v</sub> antibody construct is bivalent.
4. The F<sub>v</sub> antibody construct according to claim 3, wherein the peptide linker 2 has 11 to 20 amino acids.
5. The F<sub>v</sub> antibody construct according to claim 3 or 4, wherein the peptide linker 2 has the amino acid sequence (G<sub>4</sub>S)<sub>4</sub>.
6. The F<sub>v</sub> antibody construct according to claim 1 or 2, wherein the F<sub>v</sub> antibody construct is tetravalent.
7. The F<sub>v</sub> antibody construct according to claim 6, wherein the peptide linker 2 has 3 to 10 amino acids.

8. The F<sub>v</sub> antibody construct according to claim 6 or 7, wherein the peptide linker 2 comprises the amino acid sequence GGPGS.

9. The F<sub>v</sub> antibody construct according to any of claims 1 to 8, wherein the F<sub>v</sub> antibody construct is multispecific.

10. F<sub>v</sub> antibody construct according to claim 9, wherein the F<sub>v</sub> antibody construct is bispecific.

11. The F<sub>v</sub> antibody construct according to any of claims 1 to 8, wherein the F<sub>v</sub> antibody construct is monospecific.

12. A method of producing the multivalent F<sub>v</sub> antibody construct according to any of claims 1 to 11, wherein DNAs coding for the peptide linkers 1, 2 and 3 are ligated with DNAs coding for the four variable domains of an F<sub>v</sub> antibody construct such that the peptide linkers link the variable domains with one another and the resulting DNA molecule is expressed in an expression plasmid.

13. Expression plasmid coding for the multivalent F<sub>v</sub> antibody construct according to any of claims 1 to 11.

14. The expression plasmid according to claim 13, namely pDISC3x19-LL.

15. The expression plasmid according to claim 13, namely pDISC3x19-SL.

16. The expression plasmid according to claim 13, namely pPIC-DISC-LL.



17. The expression plasmid according to claim 13, namely pPIC-DISC-SL.

18. The expression plasmid according to claim 13, namely pDISC5-LL.

19. The expression plasmid according to claim 13, namely pDISC6-SL.

20. Use of the multivalent F<sub>v</sub> antibody construct according to any of claims 1 to 11 for the diagnosis and/or treatment of diseases.

21. Use according to claim 20, wherein the diseases are viral, bacterial or tumoral diseases.

FIGURE 5

6/10

EcoRI RBS PelB leader NcoI

1 GAATTCATTAAGGCGAATTAACCATGAAATACCTATTGCGTACCGGCGCCCGTGGCTTGGCTGCTGGCGAGCTCAGCGCGGCGATGG

131 M K Y L L P T A A A G L L L A A Q P A M

92 CCGAGGTGCACTGCAGCAGTCTCGGGCTGCACTGGCAAGACCTGGGGCGCTCAGTGAAGATGTCTGCAAGGCTTCTGGCTACACCTTTAC

222 A Q V Q L Q Q S G A E L A R P G A S V K M S C K A S G Y T F T

CDR-H1 Frame-H2 CDR-H2

183 TAGGTACACGATGCACCTGGGTAAACAGAGGCGCTGGACAGGGCTCTGGAATGGATTGGATACATTAATCCCTAGCGCTGGTTATAC

522 R Y T M H W V K Q R P S Q G L E W I G Y I N P S R G Y T

Frame-H3

267 TAATTACAATCAGAAGTTCAAGGACAGGCGCATTGACTACAGACAAATCTCCAGCAGAGCTACATGCAACTGAGCAGCGCTGAC

802 N Y N Q K F K D K A T L T T D K S S S T A Y M Q L S E L T

CDR-H3 Frame-H4

354 ATCTGAGGACTCTGCACTCTATTACTGTGCAAGATATTATGATGATCATTAAGCGCTTGACTACTGGGGCGAGGCGCACTCTCA

1092 S E D S A V Y Y C A R Y Y D D H Y S L D Y W G Q G T T L

CH1 Linker 1 Frame-L1 VL anti-CD19

440 CAGTCTCTCTCAGGCGAAGCAGCGGAGCTGGGGGGTATCTCTGCTACCGCAACTCCAGCTTCTTTGGCTGTGTCTCTAGGCGAGA

1332 T V S S A K T T F K L G G D D L L T Q T P A S L A V S L S Q

CDR-L1 Frame-L2

530 GGGCGCCATCTCTCTGCAAGGCGCAGCGCAAGTGTGATATGATGGTGATAGTTATTTGAAGCTGGTACCAACAGATCTCCAGGAC

1682 R A T I S C K A S Q S Y D Y D G D S Y L N W Y Q Q I P Q

CDR-L2 Frame-L3

614 AGCGACCGCAACTCTCTCTATGATGATGCAATCTAGTTTCTGGGCTCCGACCGCAGCTTTAGTGGCAGTGGCTCTGGGACAGACTT

1962 Q P P K L L I Y D A S N L V S G I P P R F S G S G S G T D F

CDR-L3 Frame-L4

702 CACCGCTCAACATCTCTCTCTGTCGAGAGGTGGATGCTGCAACCTATCTCTCTCAGCAAAGTACTGAGGATCCCTGGACGCTCTGGTGA

2252 T L N I H P V E X V D A A T Y H C Q Q S T E D P W T F G G

Ckappa NotI Linker 3 PvuII Frame-H1

790 GGCACCAAGCTGGAAATCAAAAGCGGCTGCTCTGCGCGCGTGGTGGCGCCAGGGTCCGAGGTGCACTGTCAGCAGCTCTGGGGCTGAGCT

2552 G T K L E I K R A D A A A A G G P G S Q V Q L Q Q S G A E L

VH anti-CD19 CDR-H1 Frame-H2

879 GGTGAGGCGCTGGGTCTCTCAGTGAACATTTCTTGGCAGGCTTCTGGCTATGCAATTCAGTACCTACTGGATGGAATCGGTGAAGCAGAGGC

2842 V R P G S S V K I S C K A S G Y A F S S Y W M N W V K Q R

CDR-H2

968 CTGACACAGGCTCTTGAAGTGGATTCAGAGATTGGCCCTGGAGATGGTGATTAATACTACAAATGGAAGTTCAAGGGTAAAGCC

3142 P G Q G L E W I G Q I W P G D G D T N Y N G K F K G K A

Frame-H3

1051 ACTCTGACTGCAGACCAATCTCTCGCAGCAGCTTACATGCAACTCAGCAGCTTAGCATCTGAGGACTCTCGCGTCTATTCTGTGCAAGAG

3422 T L T A D E S S S S T A Y M Q L S S L A S E D S A V Y F C A R

CDR-H3 Frame-H4 CH1

1142 GGGAGACTACGACGGTAGGGCGCTTATTACTATGCTATGGACTACTGGGTCAGGCAACCTCAGTCACCGTCTCTCTCAAGCAAA

3722 R E T T T V G R Y Y Y A M D Y W G Q G T S V T V S S A K

Linker 1 Frame-L1 VL anti-CD3

1226 CAACACCGAGCTTGGCGGTGATCTCTCTCTACTGCTCTCCAGCAATCATCTCTGCACTCTCCAGGGGAGAGGTCACCTGACCTGCA

4002 T P K L G G D I V L T Q S P A I M S A S P G E K V T M T C

CDR-L1 Frame-L2 CDR-L2

1316 GTGCCAGCTCAAGTGTAAAGTTACATGAAGCTGTACAGCAGAAGTCAGGCACCTCCCGCAAAAGATGGATTATGACACATCCAA

4302 S A S S S V S Y M N W Y Q Q K S G T S P K R W I Y D T S X

Frame-L3

1401 ACTGGCTTCTGAGTCCCTGCTCACTTCAAGGGCAGTGGGTCTGGGACCTCTTACTCTCTCAATCAGCGGCTAGGCGCTGAGATGC

4582 L A S G V P A H F R G S G S G T S Y S L T I S G M E A E D A

CDR-L3 Frame-L4 Ckappa

1491 TGGCACTTATTACTGCCAGCAGTGGAGTAGTAACCCATTTCAGCTTCCGCTCGGGACAAAGTTGGAATAAAGCGGGCTGATCTGCG

4882 A T Y Y C Q Q W S S N P F T F G S G T K L E I N R A D T A

c-myc epitope His6 tail XbaI

1578 ACCCACTGGATCCGAACAAAAGCTGATCTCAGAAGAAGACCTAAACTCTCATCCATCCCAATCACTAATCTAGA

5172 P T G S E Q K L I S E E D L N S H H H H H H

FIGURE 6

7/10

941 ATGAGATTTCCTTCAATTTTACTGCTGTTTTATTTCGCAGCATCCTCCGCAATTAGCTGCTCCAGTCAACACTAC  
1 M R F P S I F T A V L F A A S S A L A A P V N T T

alpha-factor signal

1015 AACAGAAGATGAAACGGGCACAAATTCGGGCTGAAGCTGTCATCGGTACTCAGATTTAGAAGGGGATTTGATG  
25 T E D E T A Q I P A E A V I G Y S D L E G D F D

1089 TTGCTGTTTTGCCATTTTCCAACAGCACAATAACGGGTTATTGTTTATAAATACTACTATTGCCAGCATTGCT  
50 V A V L P F S N S T N N G L L F I N T T I A S I A

XhoI EcoRI

1163 GCTAAAGAAGAAGGGGTATCTCTCGAGAAAAGAGAGGCTGAAGCTCAATTCCAGGTGCAACTGCAGCAGTC  
75 A K E E G V S L E K R E A E A E F Q V Q L Q Q S

VH anti-CD3

1234 TGGGGCTGAACTGGCAAGACCTGGGGCCTCAGTGAAGATGTCCTGCAAGGCTTCT  
98 G A E L A R P G A S V K M S C K A S

FIGURE 7

8/10

941 ATGAGATTTCCTTCAATTTTACTGCTGTTTTATTCCGAGCATCCTCCGCATTAGCTGCTCCAGTCAACACTAC  
 1▶ M R F P S I F T A V L F A A S S A L A A P V N T T

alpha-factor signal

1015 AACAGAAGATGAAACGGGCACAAATTCGGGCTGAAGCTGTCATCGGTTACTCAGATTTAGAAGGGCATTTCGATG  
 25▶ T E D E T A Q I P A E A V I G Y S D L E G D F D

BsrDI

1089 TTGCTGTTTTGCCATTTTCCAACAGCACAAATAACGGGTTATTGTTTTATAAATACTACTATTGCCAGCATTTGCT  
 50▶ V A V L P F S N S T N N G L L F I N T T I A S I A

XhoI EcoRI

1163 GCTAAAGAAGAAGGGGTATCTCTCGAGAAAAGACAGGCTGAAGCTGATTCATGGCGCAGGTGCAACTGCAG  
 75▶ A K E E G V S L E K R E A E A E F M A Q V Q L Q

VH anti-CD3

1235 CAGTCTGGGGCTGAACTGGCAAGACCTGGGGCCTCAGTGAAGATGTCCTGCAAGGCTTCT  
 99▶ Q S G A E L A R P G A S V K M S C K A S

FIGURE 8

UNSCANNABLE ITEM

RECEIVED WITH THIS APPLICATION

(ITEM ON THE 10TH FLOOR ZONE 5 IN THE FILE PREPARATION SECTION)

DOCUMENT REÇU AVEC CETTE DEMANDE

NE POUVANT ÊTRE BALAYÉ

(DOCUMENT AU 10 IÈME ÉTAGE AIRE 5 DANS LA SECTION DE LA  
PRÉPARATION DES DOSSIERS)

P1-2-3-4-9-10